U.S.S.N. 09/768,155
Filed: January 23, 2001
CLEAN UP VERSION OF AMENDED SPECIFICATION PARAGRAPHS PURSUANT TO 37 C.F.R. § 1.121 (b)(1)(ii)

Clean Version of Amended Specification Paragraphs

Pursuant to 37 C.F.R. § 1.121(b)(1)(ii)

-- Figures 2a and 2b are the nucleotide (SEQ ID NO:1) and deduced amino acid (SEQ ID NO:2) sequences of the cDNA insert in G7. Asterisks show the stop codon. This sequence has been added to the GenBank nucleic acid sequence database, Los Alamos National Laboratory, NM, and has been assigned accession number U27517. --

-- Figure 3a and 3b are an alignment of the central core regions of 5 ribosomal proteins (SEQ ID NOS:3 to 7)(r-proteins) S1. Asterisks show the 5 repeating regions (SEQ ID NOS:8 to 12)(a, b, c, d, and e, respectively). Spaces indicate positions where gaps were introduced to optimize alignment of the sequences (SEQ ID NOS:3 to 7). Dashes indicate identity to the residues of HS1 (SEQ ID NO:3). Alignment of the central core region of HS1 (SEQ ID NO:3) is residues 63-317. HS1; human r-protein S1 (SEQ ID NO:3)presented in this study, ES1; E. coli r-protein S1 (SEQ ID NO:4)(Ref. 26), RS1; Rhizobium melilotii r-protein S1 (SEQ ID NO:5)(Ref. 28), PS1, Providencia sp. r-protein S1 (SEQ ID NO:6)(Ref. 27), CS1; chloroplast r-protein S1 (SEQ ID NO:7)(Ref. 29).--

-- Although described herein with reference to the whole protein, it is preferable to use peptides of between a few amino acids up to about 100 amino acids, more preferably less than forty amino acids, still more preferably less than ten to twenty amino acids. These peptides can be easily ascertained by immobilizing the anti-dsDNA antibodies from a patient(s) and screening for binding of the peptides. Peptides can be prepared using standard techniques for amino acid synthesis or recombinantly, by engineering the cDNA (SEQ ID NO:1) encoding the protein, described in Figures 2a and 2b.--

U.S.S.N. 09/768,155
Filed: January 23, 2001
CLEAN UP VERSION OF AMENDED SPECIFICATION PARAGRAPHS PURSUANT TO 37 C.F.R. § 1.121 (b)(1)(ii)

-- The nucleotide sequence of the cDNA insert in G7 was determined. Its primary nucleotide (SEQ ID NO:1)and deduced amino acid (SEQ ID NO:2) sequences (GenBank no. U27517) are shown in Figures 2a and 2b. The cDNA insert (SEQ ID NO:1) proved to be 1,314 nucleotides in length. The TAA stop codon is located at positions 1057-1059. The predicted molecular weight for the encoded polypeptides (SEQ ID NO:2) (352 amino acids) is 38.0 kDa. However, this cDNA insert (SEQ ID NO:1) in G7 seems to be a partial length cDNA because the molecular weight of the encoded polypeptide (SEQ ID NO:2) is smaller than the estimated full length size (104 kDa) of the reactive protein in MOLT4 cell extract. Thus, this cDNA (SEQ ID NO:1) does not seem to contain the initiation codon.--

-- A search for similarities between the nucleotide sequence of the cDNA (SEQ ID NO:1) in G7 (GenBank no. U27517) and other sequences through the NCBI using the BLAST network service showed a significant match (99% identity) with a sequence encoding human ribosomal protein (r-protein) S1 homologue mRNA reported by Eklund at al., *Gene* 155:231 (1995) (SEQ ID NO:3). However, there are 3 nucleotide and 1 amino acid differences between the G7 cDNA insert(SEQ ID NO:1) and their cDNA sequence (SEQ ID NO:3) (GTC (positions 130-132) in the G7 cDNA insert (SEQ ID NO:1) vs GTA (positions 292-294) in their cDNA (SEQ ID NO:3), AGT (positions 133-135, encodes Ser at residue 45) in the G7 cDNA insert (SEQ ID NO:1). vs GCT (positions 295-297, encodes Ala at residue 99) in their cDNA (SEQ ID NO:3)]. Moreover, 2 nucleotides (C at positions 1355 and 1366) and 162 nucleotides (positions 1-162) in their cDNA (SEQ ID NO:3) are deleted in the G7 cDNA insert (SEQ ID NO:1). A search was made for some similarities between the predicted amino acid sequence (SEQ ID NO:2) and other protein sequences in the SWISSPROT database using the algorithm as described by

Gish, et al., *Nature Genetics* 3:266 (1993); Altschul, et al., *J. Mol. Biol.* 215:403 (1990). --

-- Identify and similarity with r-proteins S1 are the following; 39% identity and 65% similarity with E. coli r-protein S1 (ES1) (SEQ ID NO:4) (26), 40% identity and 64% similarity with Providencia sp. r-protein S1 (PS1) (SEQ ID NO:6) (Schnier, et al., *Mol. Gen. Genet.* 200:476 (1985)), 38% identity and 63% similarity with Rhizobium melilotti r-protein S1 (RS1) (SEQ ID NO:5) (Schnier, et al., *Nucleic Acids Res.* 16:3075 (1988)), and 50% identity and 71% similarity with chloroplast r-protein S1 (CS1) (SEQ ID NO:7) (Franzetti, et al., *J. Biol. Chem.* 267:19075 (1992)). Moreover, 5 repeating regions [EGTV (SEQ ID NO:8) due 158-161 and 243-246), DFGAFV (SEQ ID NO:9) (166-171 and 251-256), GLVHVS (SEQ ID NO:10) (178-183 and 264-269), GDKV (SEQ ID NO:11) (200-203 and 286-289), and RISLS (SEQ ID NO:12) (216-220 and 302-306)] were observed in the protein sequence (SEQ ID NO:2). These repeating residues (SEQ ID NOS:8 TO 12) have a high degree of homology among other r-proteins S1 (Figures 3a and 3b).--

-- In summary, a lambda gt11 cDNA library constructed from mRNA of human liver was screened by using a SLE patient serum with anti-dsDNA antibody and a clone G7 which has a 1.3-kb cDNA insert (SEQ ID NO:1) isolated. Not only all of the 10 anti-dsDNA patient sera but also affinity-purified anti-dsDNA and human IgG monoclonal anti-dsDNA antibodies recognized the protein expressed by G7. The affinity-purified antibody eluted from this protein was positive for anti-dsDNA antibody activity by the Crithidia assay. Moreover, antibody binding to this protein was inhibited completely by DNA but not by RNA. From those observations, it was concluded that anti-dsDNA antibodies cross-react with the protein expressed by G7.--

-- A significant match (99% identify) between the nucleotide sequence of the cDNA in G7 (SEQ ID NO:1) and a cDNA (SEQ ID NO:3) reported by Eklund et al. as encoding human r-protein S1 homologue mRNA was found. It appears that anti-dsDNA antibodies directly bind to the protein expressed by G7 because the reactivity of anti-dsDNA antibodies against the protein was not influenced by DNAase I treatment and the binding of anti-dsDNA antibodies to the protein was inhibited completely by DNA. The predicted amino acid sequence (SEQ ID NO:2) presented in this study had homology with some r-proteins S1 including ES1 (SEQ ID NO:4). ES1 (SEQ ID NO:4) is well characterized at the functional and structural level (Subramanian, *Prog. Nucleic Acids Res. Mol. Biol.* 28:101 (1983)) while there are few reports about mammalian r-proteins S1. ES1 (SEQ ID NO:4) is the largest protein of the ribosome and has the same length as the ribosome. This protein is associated with the 30S ribosomal subunit in prokaryotes via its N-terminal globular domain and is known to stimulate translation by facilitating mRNA binding to the 30S ribosomal subunit. The central and C-terminal region contain repeating homologous sequences which are known to play a key role in the binding of structural elements of r-protein S1 to mRNA. Five repeating regions (EGTV (SEQ ID NO:8), DFGAFV (SEQ ID NO:9), GLVHVS (SEQ ID NO:10), GDKV (SEQ ID NO:11) and RISLS (SEQ ID NO:12)) which repeat twice in the central core region of the protein (SEQ ID NO:2) were observed (Figures 3a and 3b). This apparent gene duplication which encodes this repeat region is absent only in the chloroplast r-protein S1.--

-- It has been reported that anti-dsDNA antibodies have high frequencies of basic amino acids carrying positive charges in the heavy chain complementarity determining regions and that arginine is the most versatile amino acid for binding

with negative-charged DNA (31). However, there are no high scoring negative charged segments, which could be an epitope for anti-dsDNA antibody, in the primary sequence of HS1. These observations might suggest that cross-reactions between anti-dsDNA antibody and HS1 are not dependent on charge interaction in the primary sequence alone but rather that the cross-reactive epitope depends on conformational apposition of negative charges in the tertiary structure of HS1. However, it is likely that HS1 mimics DNA because anti-dsDNA antibodies cross-react with HS1. It is also appealing to believe that proteins which "mimic" the structure of DNA could play a role as immunogen.--

-- It is assumed that most of the amino acid sequences which G7 encodes are mRNA-binding sites on HS1 because five repeating regions (SEQ ID NOS:8 to 12) (residues 158-306) which repeat twice in HS1 (SEQ ID NO:3) and are supposed to be a feature of the mRNA binding site are observed. Therefore, if anti-dsDNA antibodies like 33.H11 bind to mRNA-binding portions on HS1, these anti-dsDNA antibodies might block the binding of mRNA to 40S (eukaryotes) ribosomal subunit, that is, the initiation reaction of translation.--